DOCUMENT RESUME

ED 106 502 CE 003 668

AUTHOR Altschuld, James W.; And Others
TITLE Occupational Exploration Program: Evaluation Report Notebook--An Overview.
INSTITUTION Ohio State Univ., Columbus. Center for Vocational and Technical Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Aug 74
NOTE 42p.; For individual reports see CE 003 669-75

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS *Career Education; Educational Assessment; Educational Objectives; *Evaluation Methods; Formative Evaluation; Junior High Schools; *Measurement Instruments; Measurement Techniques; Pilot Projects; *Program Descriptions; Program Design; *Program Evaluation; Reports; Simulation; Summative Evaluation
IDENTIFIERS *Career Exploration; Occupational Exploration Program; OEP

ABSTRACT

The overview serves as an introduction to seven evaluation reports produced for the Occupational Exploration Program (OEP), a joint developmental effort of the Center for Vocational Education, Ohio State University, and the Jefferson County, Colorado public schools. The 1973-74 project was designed to provide occupational exploration experiences for junior high school students. The report contains a description of evaluation procedures employed by the program as well as an overall description of the program itself. The latter includes goals, objectives, and a discussion of simulation, the major technique used for delivering the program. The evaluation procedures section covers summative and formative evaluation concerns; general sampling; design parameters; and a brief description of the Reviser's Information Summary, the compilation of evaluation data that accompanies each of the seven OEP reports. The final section describes the evaluation instruments utilized and procedures followed. Discussed are: knowledge tests, affective tests, student module questionnaires, teacher evaluation logs, panel review procedures, and observer forms. (Author/MW)

ED106502

FEB 28 1975

OCCUPATIONAL EXPLORATION PROGRAM

EVALUATION REPORT NOTEBOOK - AN OVERVIEW

Prepared By

James W. Altschuld

With The Assistance of

Norman Singer
Brian Fitch
Arthur Terry
Robert Klabenes
Roger Brown
Sandra Pritz
Robert Campbell
Robert Cotman
Jackie Lechner
Robert Blum (Jefferson County, Colorado)
John Radloff (Jefferson County, Colorado)

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Patricia B. Hannen
Product Utilization, CVE
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

THE CENTER FOR VOCATIONAL EDUCATION
THE OHIO STATE UNIVERSITY

CE003648

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

August 1974

The project presented/reported herein was performed pursuant to a grant from the National Institute of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

Copyright 1974, by the Ohio State University, The Center for Vocational Education.

Copyright for these materials is claimed only during the period of development, test, and evaluation, unless authorization is granted by the National Institute of Education to claim copyright also on the final materials. For information on the status of the copyright claim, contact either the copyright proprietor or the National Institute of Education.

ED106502

FEB 26 1975

ABSTRACT

Evaluation Report Notebook - An Overview

Report for the Occupational Exploration Program

By James W. Altschuld

This overview is to serve as an introduction to seven evaluation reports produced for the Occupational Exploration Program. The Occupational Exploration Program (O.E.P.) is funded by the National Institute of Education (N.I.E.) and is a joint developmental effort of the Center for Vocational Education (The Ohio State University) and the Jefferson County, Colorado public schools. The overview contains a description of the evaluation procedures employed for the program as well as an overall description of the program itself. The latter includes goals, objectives and a discussion of simulation, the major technique used for delivering the program. The evaluation procedures section covers summative and formative evaluation concerns; general sampling and design parameters; instrumentation; and a brief description of the Revisor's Information Summary, a unique way of compiling information collected in the evaluation.

IE003668

ACKNOWLEDGEMENTS

Evaluation endeavors are usually the result of the efforts of many individuals. Therefore the author wishes to thank:

1.) The teachers, administrators, and students in Jefferson County, Colorado and Denver, Colorado who participated in the use and evaluation of Occupational Exploration Materials;

2.) Jon Schaffarzick, Michael Hock, and David Hampson of the National Institute of Education for their support of the Occupational Exploration Program; and

3.) The eleven staff members identified on the cover, who, by their support, expertise, and/or direction, helped to make the evaluation of Occupational Exploration Materials possible.

TABLE OF CONTENTS

		PAGE NUMBER
I.	INTRODUCTION	
	A. Purposes/Description of the Notebook	1
	B. How to Use the Notebook	2
II.	PROJECT DESCRIPTION	
	A. Overview and Types of Product	4
	B. Goals/Objectives	6
	C. Intended Audience and Use for the Simulations	9
III.	EVALUATION PROCEDURES - DESCRIPTION OF	
	A. Overview	10
	B. Types of Evaluation	11
	C. Sampling	15
IV.	INSTRUMENTATION	
	A. Knowledge Test (What Do You Know?)	20
	B. Affective Test (What Do You Like?)	23
	C. Student Module Questionnaires (What Do You Think?)	26
	D. Teacher Evaluation Log	29
	E. Panel Review Procedure	32
	F. Observer Form	34
V.	SUMMARY - FURTHER CONSIDERATIONS	35

I. Introduction

A. Purposes/Description of the Notebook

The contents of the evaluation report notebook represent a compilation of the evaluation data collected during the pilot test of occupational exploration curriculum modules developed jointly by the Center for Vocational Education and the Jefferson County, Colorado Public Schools. The notebook is organized as follows:

- There is a front section in which overviews of the general nature of the project and the evaluation that was conducted are given. This introduction is part of that front material.
- The next sections of the notebook contain specific module evaluation reports including a description of the actual evaluation that was implemented for a module, data tables and interpretation, and of greatest importance, a "Reviser's Information Summary". (Note: Module specific sections will contain outlines describing their individual content.)

The notebook has been designed to serve several basic purposes. First, it will be a major record source for project personnel regarding pilot test data collected for the project in 'FY' 73-74. The scope of the evaluation and the amount of data being collected necessitate that the data be summarized, collated and easily accessible for future use. The notebook will facilitate the process of accessing summarized data. Secondly, it will provide an overview of the project evaluation for both project personnel as well as the project sponsor. Lastly, the Reviser's Information Summary (RIS) provides the means for module revisers to start the revision process without having to wade through an excessive amount of data. If the reviser does desire to study data the notebook also includes data tables for his/her perusal.

B. How to Use the Notebook

1. For Revisers

On initial use of the notebook it is recommended that the reviser first read the front section. This will enable him to understand the overall nature of the evaluation strategy employed on the project. Next the reviser should review the module specific section which contains descriptions of how the evaluation was actually carried out, sample used, type of test questions, etc., and skim Section III, Parts A, B, C and D. Section E probably should be read carefully. This review will help the reviser to understand the context in which the module was used and should help in interpreting the Reviser's Information Summary. Then he should proceed to the Reviser's Information Summary for the specific module and closely read the information provided. This coupled with his/her knowledge of the module should enable the reviser to begin to generate a revision strategy for the specific module. If the reviser feels the need to examine the data more closely, the next step would be to refer back to the appropriate data tables provided in the report. In other words, after reading the RIS the reviser may use the data tables in accordance with his/her individual information needs. When using the notebook for a second time the reviser most likely should omit reading this general front matter portion of the notebook.

2. For Evaluators, Project Supervisory Personnel, NIE Staff

As in the case of revisers, it is recommended that evaluators, program supervisory personnel and NIE staff read the general front matter first and then read the entire module specific section. On second use of the notebook, it would then only be necessary to read the module specific section.

3. Special Note

Throughout this front material, commentary, as already found in the text, has been added to assist the reviser in developing a perspective on the revision process. These comments are by design just a sample of brief suggestions or things to consider as the reviser contemplates the revision process.

II. Project Description

A. Overview and Types of Product

The occupational exploration project was designed for the purpose of assisting students in the middle grades (8th and 9th grades) with the process of exploring occupations. It was to not only help them explore occupations, but also to aid them in developing future occupational and/or occupational exploration plans. The basic strategy for working with students was to produce curriculum materials that could be implemented with small groups or classes of students. The instructional technique for most of the materials produced in 1973-74 centered around simulations that would take approximately 3-4 weeks of class time (1 hour per day). The simulations were built with the following considerations in mind:

- That they would expose students to a variety of careers in the occupational area being simulated (Breadth of Coverage);
- That they would provide the students with the opportunity to study at least one career in some depth (Depth of Coverage);
- That they would provide "hands on" or active types of roles for the students to play as opposed to a passive or a "telling" students type of approach (Hands On Experience);
- That they would be heavily student directed, not self instructional but student directed. The teacher's role is more that of a facilitator, guider or leader -- a manager of instruction as opposed to a director of instruction (Student Directed);
- That they would include mediated alternatives (slides, cassettes, films, talking pages, etc.) at key points within the simulation so that students would not always have to read the materials.

It was hoped that this approach would enhance students' motivation as they worked their way through the materials (Mediated Motivation).

For the project year 1973-74, 12 simulation modules were to be developed, with seven (7) of those modules to undergo extensive pilot testing in the Jefferson County, Colorado, and Denver, Colorado public schools. Where feasible, incremental testing (testing of module parts) was to be carried out during the development of modules. The developmental, as well as the testing schedule, is briefly outlined in Table I.

Table I - Schedule of Development/Testing for Modules

Time	Development	Testing
Autumn 1973	3 simulation modules	Pilot Test
Late Autumn, 1973-Late Winter, 1974	4 simulation modules	Pilot Test, Incrementally Test Where Feasible
Winter - June 30, 1974	5 simulation modules (Up to Pilot Test Readiness)	Incrementally Test, Where Feasible

A simulation module consists of four main parts. The first part, the Preview, consists of a short introductory section - usually a film or sound-slide presentation with an accompanying illustrated booklet. Its primary function is to motivate or stimulate student interest in participating in the simulation. In the ideal world, the individual student would elect to either participate or not participate after seeing/reading the preview. (As some schools are presently constituted, this may not be feasible and students

generally might not be allowed to exercise the option). After seeing/reading the preview, the students move into the Preparation Phase which enables them to identify the role(s) they will perform in the simulation. The students then enter into the activity or Participation Phase of the simulation. Here they perform various tasks which are related to their roles and help them to better understand the nature of the role as it might occur in the World of Work. Upon completion of the participation phase, the students move into the Summary Phase of the module. In the summary phase, students analyze and pull together their exploration experiences in the module. Hopefully, this will help them plan and prepare for future exploration experiences both of a formal, as well as an informal nature.

The parts of a module are especially important to the module reviser inasmuch as the revision must deal with them specifically. In addition, the reviser must pay attention to the transition points between the phases. Inadequate transitions may seriously effect the success of the module. Data will be summarized, to the extent possible, on a phase by phase basis to facilitate the revision process. The reviser should keep in mind that only a subset of potential revision questions could be answered at any one time. Care therefore should be taken in reading the module and studying the revision information before the development of a revision strategy is undertaken.

B. Goals/Objectives

As stated earlier the basic purpose of the project was to help students explore careers and develop career exploration plans. More specifically a set of project goals and sub-goals were developed. These are listed below (Table II). The reader should keep in mind that while these goals provided the focus for the development of modules, a much more specific set of goals - objectives is necessary for conducting the evaluation. As a

first step at establishing the objectives, CVE and Jeffco project staff took each sub-goal and delineated behaviors that might be representative of the goal. This effort resulted in an extensive listing of behaviors - one which was fully beyond the present capability to evaluate. A further consideration was that some of the delineated behaviors, as well as the original goals, were more closely related to the overall impact of a large scale program as opposed to the intervention of one or two modules. Therefore, project staff decided upon a more basic approach to the interim assessment of module effects or impact - that is, a simplified set of behaviors was established for the pilot test of seven (7) modules produced in 1973-74. These behaviors, with the limitations given in the next paragraph were measured in the interim summative evaluation of a module.

This set of behaviors is listed in Table III. Note that these behaviors were generated from an overall project point of view. A given module will not necessarily deliver equally on all of the behaviors and for that matter, a given module might deliver on only two or three of the set of behaviors. Specific module intents and contents do differ within the general framework of the project. In each module specific section, a description of the behaviors delivered by that module will be given.

Table II - Goals and Sub-Goals as Specified in the Project Proposal (1973-74)

1.1 Sub-Goals Related to the Development of an Information Base

1.1.1 To build an occupational information base regarding the environment, the people, and the processes associated with the world of work.

1.1.2 To build a personal information base regarding interests, aptitudes, achievements, and values associated with the world of work.

1.1.3 To develop increased awareness and recognition of the ways in which people and occupations within a broad field of work and across several broad fields of work resemble or differ from one another.

1.1.4 To apply several occupational classification systems in characterizing broad fields of work.

1.2 Sub-Goals Related to the Development of Attitudes

1.2.1 To heighten one's sensitivity toward perceiving and responding to experiences associated with the world of work.

1.2.2 To develop personal feelings and preferences toward occupational exploration experiences.

1.2.3 To develop an appreciation for potential personal fulfillments associated with different occupational opportunities.

1.3 Sub-Goals Related to the Development of Applicative Strategies

1.3.1 To formulate and refine plans to explore in more relevant detail broad fields of work based on a synthesis of occupational and personal information.

1.3.2 To systematically apply exploration techniques to personally

appropriate occupational decisions.

1.3.3 To compare hypotheses regarding one's developing self and occupational preferences with one's past preferences and stages of development.

1.3.4 To generate and use a set of criteria to evaluate the qualitative realism and goodness of personally derived occupational exploration goals.

Table III - Delineated Behaviors for Use in the Pilot Test (1973-74)

AREA	TYPE OF BEHAVIOR
Cognitive Skills	Knowledge of: Processes Responsibilities Tools Environments Skills/Requirements Interactions
	Applications of the above areas where feasible
Affective Skills	Expressed likes/dislikes (preferences for) jobs or functions in simulations as well as reasons for same.

C. Intended Audience and Use for the Simulations

1. Audience - The target audience (for 'FY' 74) modules is 8th and 9th graders from so called "average" junior high schools. Average in this case is loosely defined as schools which basically draw their populations from communities with a middle class socio-economic-status. The materials as now constituted are not designed particularly for use in an inner city school.

As for subject matter, the modules could flexibly be used in a variety of academic classrooms. Several modules are exceptions to the rule, but in

general it holds.

2. Use - The modules usually accommodate 10-15 students. Again the rule is general and several modules include roles for all the students in a class. There are four or five basic ways of obtaining or setting up groupings of students for a module. These will be described under the sampling part of this front section.

III. Evaluation Procedures - Description of

A. Overview

A primary factor in designing the evaluation strategy for the Occupational Exploration Program was that materials developed during the first year of the project would be in semi-finished but not final draft form. Hence, any testing or evaluation should focus primarily on the collection of revision data, i.e., the evaluation should be formative in nature. Beyond the formative emphasis it seemed sensible to collect information regarding the interim effect/impact of the materials on students. Therefore, a secondary emphasis of the evaluation was a summative one. The distinction just made is to a degree arbitrary and semantic, but at the same time it serves a useful purpose. That is, it provides a framework in which to describe the evaluation strategy.

In addition to the above distinction the reader must note that the pilot test of each module was viewed as an individual experiment. A full exploration program will not be in operation until approximately 2-3 years from the date of project inception. The resultant effect is that in the cognitive and affective areas the assessment of only a limited set of objectives (as shown earlier in Table III) could be undertaken.

B. Types of Evaluation

1. Formative Evaluation

a. Purpose

The overall goal of formative evaluation for the Occupational Exploration Program in 1973-74 is to develop a bank of information relevant and useful for revision and refinement of each module that is pilot tested. More specifically the formative evaluation will attempt to collect information regarding: problems in implementing the modules; suggestions regarding the improvement of the module; and overall perceptions of the module.

b. General Design and Information Sources

Data for the formative evaluation was collected from two groups - teachers and students who participated in the module implementation. The basic approach was for teachers to complete sections of a questionnaire, the Teacher Evaluation Log, as they worked with and observed students progressing through a module. The questionnaire was designed and ordered so that each point in it correlated with a specific part of the module. The last section of the Teacher Evaluation Log contained questions regarding the teacher's overall perception of the module. Besides completing the LOG, teachers also participated as a group in a post module debriefing session (Post Module Panel Review). Here they could freely* discuss what they considered to be the strengths and weaknesses of the module and the

*Note: The panel always did have a leader (a Jeff Co. Administrator) who helped to keep the general discussion of a module on track and geared toward the eventual revision of a module. CVE and Jeff Co. personnel worked closely together to establish guidelines for conducting the panels.

types of changes that they would make in it. The results of the panels were recorded, and summarized for use in the revision process.

For students, revision data was collected by means of questionnaires administered to students after they had completed a module and from pre- and posttests dealing with the aforementioned cognitive and affective behaviors. This latter set of data is limited from a revision standpoint, however, for two reasons. First, the tests were developed from an overall module standpoint and it is difficult to pinpoint the data directly to a specific module section - a key concern for revisers. Secondly, and related to the first concern, is that the tests were written for the purpose of assessing module effect/impact, not for diagnostic purposes. The reviser is, therefore, cautioned about placing undue emphasis upon the results from the cognitive and affective tests.

There are several other minor sources of revision data. These include observations of Jeffco project staff, reviews of materials as they were developed, comments of editors and product developers, etc. Where these are available, the revision information will be included on the Reviser's Information Summary.

In Table IV a brief overview of the formative evaluation portion of the pilot test is shown.

Table IV - Overview* of the Formative Evaluation Portion of the Pilot Test (1973-74)

Concern	Information Source: Teachers	Information Source: Students	Timing Relative to When Module is Taught: Before	During	After	Design	Instrumentation--
1. Revisions Suggestions	X**	X		X**	X	Survey (Descriptive Techniques	Teacher Questionnaire(s) Student Questionnaire Panel Reviews On-Site Observations ***
2. Implementation Problems	X	X		X	X		
3. Overall Perceptions	X	X			X		

* The table does not include minor sources of revision data as explained in the text.

** An X in a column indicates that the data is collected from a particular source or the timing of when the data is collected, respectively.

*** Applies only to the 4 Modules tested in the Spring of 1974.

2. Interim Summative Evaluation

a. Purpose

The purpose of the interim summative evaluation was to collect data regarding the impact of the modules on students and to provide an additional source of data input into the revision process. Information concerning the amount and type of student gain on cognitive and affective measures was sought.

b. General Design and Sources of Information

The basic design* of the evaluation can be generically diagrammed as:

R	O_1	X_1	O_2	Experimental Group
R	O_3		O_4	Control Group

Wherein O_1 and O_3 represent pretests, O_2 and O_4 represent posttests, X_1 stands for the experimental treatment (i.e., the module) and R equals random assignment of individuals to treatment group. Stated differently, the design consists of comparable experimental and control groups, where the experimental group received the module and the control group did not. Both groups would receive two (2) instruments as a pretest and the same two as a posttest. The instruments consisted of a multiple choice type knowledge test and a brief attitude scale with some additional open-ended questions. The questionnaires and data sources described earlier for formative evaluation also contain opinion type questions which will help to determine the impact of the modules.

---

*A slight variant of this design was used for the second set of 4 modules tested in the Spring of 1974.

Data collected from the two groups will be presented in tables in the module specific sections of the notebook. For the most part, the tables will be descriptive in nature. Some of the data, however, will be analyzed through a rather complex statistical procedure. To assist both the reviser and the reader of the notebook data tables along with brief interpretations of the data will be included in the module specific section.

Table V is a descriptive overview of the Interim Summative Evaluation Portion of the Pilot Test.

C. Sampling

1. General Sampling Framework

Ideally, the sampling process would be based on a randomization or a randomized procedure. By means of randomization (either assignment or selection) it would be possible to say that both the control and experimental groups were, within the laws of chance, equal. That is, the groups would essentially be equalized in terms of I.Q., reading ability, prior knowledge of careers, etc. Furthermore, if students were randomly selected and/or assigned to control and experimental groups from a well defined population it would be possible and legitimate to generalize the results of the evaluation back to that population.

Given the above framework, the difficulties of sampling in this situation become apparent. First, to work in the schools with modules of the type produced in this project requires the cooperation of many school personnel. Small groups making extensive use of media will in many instances change or alter normal school operation and regular classroom routine. As a

Table V - Overview of the Interim Summative Evaluation Portion of the Pilot Test

Concern	Information Source		Timing Relative to When Module is Taught			Design	Instrumentation
	Teachers	Students	Before	During	After		
1. Cognitive Impact		X*	X*		X	Pretest-Post-test Control Group Design	Measure of Cognitive Achievement per module (Paper and Pencil)
2. Affective Impact		X	X		X		Measure of Affective Achievement per Module (Paper & Pencil)
3. Overall Perceptions	X	X			X	Survey (Descriptive Techniques)	Teacher Questionnaire(s) Student Questionnaire Panel Reviews
4. Estimates of Learning	X	X		X	X		
5. Description of Sample	-	-	-	-	-	-	Brief Descriptions Written by Local School Building Administrators

*An X indicates that the data is collected from a particular source or the timing of when the data is collected, respectively.

result, it may have been harder to locate schools willing to participate in the pilot test of the modules. Thus the schools and the teachers who participated in the pilot test of a module were volunteers. Volunteers may differ from the so called average school or teacher (whatever those undefined entities might be). Secondly, most schools cannot easily handle the random assignment of students to control and experimental groups - hence, most of the time, students were volunteered from the context of an intact classroom. (Note: the teacher questionnaires contained several questions regarding how students were selected for a module.)

Several other factors weighed heavily in the sampling procedure. First, the pilot test of a module was seen as an individual experiment. The systematic study of the impact of more than one module on a student was not only not undertaken, it was consciously avoided due to the early developmental nature of materials and time constraints. This made it necessary to involve a good number of both Jeffco and Denver junior high schools due to the fact that fresh, uncontaminated samples were needed for testing each module. Second, the Denver Public School System situation is unusual in the Denver is presently busing limited numbers of students for the purpose of racially balancing the schools. The cost of collecting data to describe each school population involved in the pilot test did not seem warranted, so in the module specific section only very brief description of the particular school will be given. This will be a qualitative description developed by an administrator (or administrators) familiar with the school in question.

2. Types of Classes to be Involved

As noted earlier the modules were not designed with a specific subject area in mind. It was felt that they could flexibly be used in a variety of subject matters areas such as English, social studies, mathematics, etc. The materials offer opportunities for educators in different areas to incorporate career education into their respective curricula. The modules could also be used with volunteers from clubs, study halls, etc.

For some modules, however, specialized equipment, settings, and teachers were required. Where this is the case, the text of the module specific section will include reference to the limiting nature of the module.

When students were selected to participate in a module from intact classrooms, an attempt was made to choose control groups from classrooms similar to those of the experimental groups. For example, if an English class provided the experimental group students, then another English class was designated as a control group, where possible. If the latter arrangement could not be established then, perhaps, a social studies classroom was selected.

3. Types of Grouping

Five types of experimental groupings of students were anticipated as the pilot test was started in the Denver and Jeffco schools. These groupings were a means of accommodating the program within the constraints of existing buildings and programs. Generally, they range on a continuum from most desirable to least desirable from the standpoint of the original intent of the Occupational Exploration Program. They are:

--Small group pulled out of regular classroom; most desirable in that the simulation could take place free of competing influences;

--Small groups in regular classrooms; not as desirable as the above case in that the teacher would have to simultaneously coordinate the efforts of two groups of students and develop some sort of alternative activity for those students not participating in the module. Also, there would most likely be space and noise problems;

--Two experimental groups in regular classrooms; less desirable than the previous two groupings in that group competition is introduced as well as all the problems specified in the grouping described directly above;

--Whole class involved (doubled roles); given the limited number of materials supplied with a module this method of grouping is similar in desirability to the grouping described immediately above. The press on materials could lead to serious problems in implementing a module.

--Students selected and grouped other than by use of intact classrooms; in this instance, students would be selected (or would volunteer) from a study hall, an after school club, a gym class, etc. Although this grouping is listed last, it most closely parallels the first grouping listed above.

The importance of grouping may have considerable bearing on

how the module was received by students and teachers. The reviser should be alerted to and aware of the type of grouping as he analyzes the revision information, particularly for a grouping that may have had a negative impact on the module.

IV. Instrumentation

A. Knowledge Test (What do You Know?)

1. General Structure - as described earlier, the knowledge tests used per module were structured around the objectives found on page 8 . Each module will differentially deliver on those objectives, hence the test for one module might be more heavily focussed on knowledge of responsibilities, whereas that for another might emphasize knowledge of processes. The determination of what to test for was based upon the judgment of the project evaluator and other project staff involved in developing the module. The focus was generally limited to only two or three types of knowledge due to the amount of time allowed for testing students. (See point #3 later in this section.)

The reviser should note that the module may deliver knowledge beyond that tested for with the What Do You Know? instrument. The items in the test represent only a two time sampling of a limited spectrum of achievement. Attention must be paid to the other sources of data from the pilot test with regard to additional as well as serendipitous types of learning/achievement.

2. Types of Questions

Basically the tests consist of multiple choice questions or derivatives thereof. For example, the tests contain questions which require selecting the correct answer from one of four

choices or determining whether a statement is true or false (selecting the correct choice from one of two options). There are some matching type questions in addition to the multiple choice types.

The content and difficulty does vary somewhat. First, an attempt was made to focus on concepts presented in the modules that related to understanding careers (e.g., what are the job responsibilities of various occupations within a field). Minutia or trivial detail was avoided in question content as much as possible. Secondly, some of the questions were aimed at a somewhat higher level skill than knowledge. Each test had at least several application or problem types of multiple choice questions.

An average item difficulty of .50 is the ultimate goal in test construction. This would tend to yield the most reliable test and help to overcome the limited numbers of items that were dictated by constraints (see next section) in the pilot test. Reliability estimates will be provided in the module specific sections to indicate the degree to which the goal was achieved.

3. Approximate Length - the tests do vary in length within a set of broad constraints. Time, in terms of test administration, was crucial in that it was the judgment of project staff that no more than one class period (about 40-45 minutes) could be allotted to the pretesting and posttesting, respectively. Hence the knowledge test would have to take quite a bit of time less than 35 minutes if an affective or attitudinal measure was to be administered during the same period. Test reliability would thus tend to be decreased.

Given this constraint, the tests for the first three Modules were a maximum for the first 3 modules of 25 questions in length. However, some questions had several sub-parts so that the maximum length in terms of questions and sub-parts (which are treated as separate questions) was 45 questions. For the second 4 modules tested in the Spring of 1974, the tests generally contained about 34 questions. If the reviser does decide to study the test result tables in the module specific section, it is advised that the test length be taken into account when looking at test reliability.

4. Scoring Procedures

Each question (and sub-part) is scored as though it was a single question of equal weight. No differential weighing of questions will be applied at this time. Due to limited lead time in preparing tests for pilot testing of the first three modules, it was not possible to incorporate answer sheets into the test administration sessions. Thus, the scoring procedure assumes somewhat greater importance than it would in other circumstances. First, a code notebook was prepared, then student answers were transferred to the right hand margin of the page in order to facilitate key punching. The transferral of responses was then checked and Hollerith data cards were punched and machine verified. (Note, for the second 4 modules tested in the Spring of 1974, machine scorable answer sheets were utilized.)

Data scoring and test analysis was then conducted by use of a standard program located in the computer library of the Evaluation Center of The Ohio State University. The program produced such data as a listing of test scores per individual, test means, test standard deviations, reliability coefficients,

item indices, etc. These will be found in the module specific sections of the report.

B. Affective Test (What Do You Like?)

1. General Structure (Objectives) - the What Do You Like? scale was designed to measure or begin to assess the effect of the module on student attitude. In this case, it is difficult to define attitude. Attitude towards what? And at what level or depth of the attitudinal domain should the instrument be aimed? In addition, the constraint of time (perhaps 10-12 minutes of class time for test administration) greatly precludes the use of a lengthy instrument.

A decision was made to focus the attitudinal scale on one aspect of affective behavior. That is, what are the student's preferences for either functions or types of occupations presented in the module. Would a student as a result of participating in a module, have a stronger, more defined preference for an occupation? Could a student justify or better explain his reasons for having a preference from taking the module? The attitude scale was seen as a preliminary start in the direction of measuring a student's beginning awareness of occupations. To the end of measuring awareness, a brief scale (5-10 questions) was constructed and administered to experimental group students. Besides assessing awareness several open-ended questions dealing with selecting or exploring occupational possibilities were included in the instrument. These dealt with a student's perceptions of what types of experiences a person should have before he/she enters the world of work and what an individual should take into consideration before selecting a job.

2. Type of Questions - There were several basic types of questions included in the What Do You Like? instrument. The first type contained two parts - one dealing with a preference for an occupation or occupational function and the student's reasons for expressing that preference. The student initially responded to the preference scale and then was given space to state reasons for his choice. For some of the attitude scales a second page was added which allowed the student to supply his own occupations or functions, to state his preference for them, and to then state his reasons for that preference.

The preference scale consisted of the following four scale values:

--Yes, I would like to try this (occupation or function);

--I'm uncertain about trying this (occupation or function);

--No, I would not like to try this (occupation or function);

--I do not have enough information to make a decision.

It was felt that as a result of participating in the module students would develop stronger preferences, and ones that they could more definitively support or explain. In accordance with that perspective, the scale was scored so that the stronger the preference the higher the score. The details of the scoring system will be given in point four below.

In addition to the preference questions, three-four open-ended questions regarding how one might go about preparing for an occupation or selecting an occupation were included in the What Do You Like? instrument. For the four modules tested in the Spring of 1974, only one open-ended question was used. (This question required the students

to complete a conversation regarding jobs in the particular field being simulated. Again, it was felt that as a result of module participation the students would be able to have a better idea of the job selection/preparation process.)

3. Approximate Length

As in the case of the knowledge test, length was greatly constrained by time factors. One class period was the total time available for both the pretesting or posttesting, respectively. Given the length of the knowledge test (35-45 questions and question parts, 25-35 minutes) the attitudinal measure had to be fairly brief. In general, 10-15 minutes were allotted for administering the instrument. The scale, therefore, was about 2-3 pages long with about 8-15 questions for the first 3 modules and 6-7 questions for the second set of 4 modules tested.

4. Scoring Procedures

Scoring of the What Do You Like? scale is somewhat complex. For the preference questions, the stronger the preference the higher the score. In other words, if the student stated that he was or was not interested in a career, he received a score value of three. If he was uncertain of his preference he received a score of 2 and if he possessed insufficient information on which to make a choice he received the lowest possible value of one. The number of reasons was then tallied and for each reason a score was given based upon the degree to which the reason supported or fitted the preference choice. Reasons were also coded as to type, i.e., financial, ability, enjoyment, etc.

Other open-ended questions were scored according to number of reasons supplied, the degree to which each reason fit the question, and type of reason. In no instance were more than three reasons coded. This is based on the observation that more than three reasons being given was a fairly rare occurrence.

The strategy for scoring responses and categorizing reasons was developed after a preliminary analysis of data. Randomly selected forms were studied to see if the responses (reasons or explanations) given by students fell into categories. From this type of analysis a scoring scheme and procedure was established.

During the scoring of forms considerable judgment on the part of the scorer/rater must be employed. Detailed instructions for scoring were written and adhered to as much as possible. Records of inter-rater reliability were maintained. One person rated the forms and transferred responses to IBM machine scorable answer sheets. A second rater rated a sample of the attitude scales and the degree of consistency among the raters was established.

C. Student Post Module Questionnaires (What Do You Think?)

1. General Structure

The post module questionnaire is quite different from the knowledge and attitudinal instruments. The latter two are closely related to the interim summative evaluation that was described earlier. As noted previously, these instruments have somewhat limited utility in regard to formative or revision types of concerns. The post module questionnaire, on the other hand, is designed to assess or to collect information about, the students' perceptions of the module. For example; What did the student like most about

the module?, What did he like least about it?, What types of problems did the students find in the modules?, etc.

Students were encouraged to respond freely. The questionnaire was considered to provide some valuable information for the revision process.

2. Types of Questions

The questionnaire consists of a variety of question types. The first twenty questions are of a short attitude scale requiring approximately 5-10 minutes of the student's time. Beyond this there are a series of short questions which generally require the student to indicate his perceptions about a topic by circling a choice in a scale. Space was also provided for students to supply comments and suggestions. (For the second 4 modules, the instrument was slightly modified.)

3. Approximate Length

The module questionnaire was administered sometime (usually 3 to 5 days) after the completion of the module and the posttesting. Since the questionnaire related to perceptions of the module, it was administered to only the experimental group. The attitude scale in the questionnaire was 20 questions in length followed by 10-15 open-ended questions. Total length therefore was 30 plus questions and the average administration time was about 20-25 minutes or more. The questionnaire was free of the competing influence of other instruments in that it was administered by itself.

4. Scoring Procedures

As in the case of the What Do You Like? scale, scoring of the post module questionnaire for students is fairly complex. The first

twenty questions are an attitude scale in which the most positive response per question was given the highest score value. And conversely the most negative response per question received the lowest score value. The response to these twenty questions could be summed and treated as one scale or they could be divided into subsets of questions and subscores could be totaled. Since the scale was designed to measure or assess feelings about several different attributes of the module it is most probable that the latter course of action will be followed. (Note: The number of items per subset will be small and reliability estimates will not be computed inasmuch as they will be greatly limited in utility.)

For the remainder of the questions in the What Do You Think? questionnaire the scoring was as follows:

1. If the question contained a scale, the scale value of the students response was simply recorded;
2. For questions requiring open-ended responses, the responses were collated with attention paid to frequently repeated types of responses.

Concerns regarding the quality of the questionnaire data may arise as one begins to get involved in the revision process. Since the questionnaire covers many topics it is difficult to obtain (and/or develop) measures of reliability and validity. This problem occurs not only for the student post module questionnaire but for teacher panels and questionnaires as well. Each module specific section contains data and/or discussion regarding instrument/data collection method, reliability, and validity. However, of special importance to the reviser is the idea of information source

convergence or corroboration. While it may be difficult to ascertain if an individual piece of data is reliable/valid, if two-three independent or somewhat independent sources converge on the same point with the same information then the set of data is quite powerful. Strength emerges from the set rather than the individual data source. In other words, the trend of the data is a key point to consider in revising a module.

D. Teacher Evaluation Log

1. General Structure

The Teacher Evaluation Log differed slightly for the first two modules (Education and Construction) pilot tested and the third Manufacturing). From the standpoint of simplicity it is perhaps easiest to explain the Log for the third and then to briefly describe the difference from the first two. The Log is 6 separate instruments under one cover. The first five instruments parallel the parts of a module. That is, the first deals with the Introduction to Simulation booklet and sound-on-slide presentation that students see before starting a module. The next instrument contains questions about the module preview, the next one about the preparation phase, and so on. The last or sixth instrument relates to the teacher's overall perceptions about the module.

The Log therefore was a device for teachers to record their comments as well as their answers to specific questions as they worked with and observed students progressing through a module. The Log differed in the case of the first two modules in that the last instrument, the General Module Evaluation for teachers, was

separated from the other five instruments. The results from the Log may have been affected due to the fact that some teachers felt the quality of the last instrument was somewhat better than the first five.

Teachers were not remunerated for the time necessary to complete evaluation instruments. (Note: they were paid for the time necessary to train them for using the modules and for participating in the post module panel reviews.) This coupled with the length of the Log led to spotty completion of the forms. The reviser should be aware of this occurrence as he works with the revision data.

For the 4 modules tested in the Spring of 1974, the Log was deleted from use and a mid-module and post module teacher questionnaire were substituted in its stead. In addition, optional forms - a Media Checklist and a Daily Inventory of Perceptions (DIP) were made available to teachers. These last two forms allow teachers to comment specifically on the media used and to provide project staff with any other opinions they might have, respectively.

2. Types of Questions

Question form or type varied considerably. Three basic types of questions will be found in the Log and in the teacher questionnaires. They are:

1. Scaled questions with space for open-ended comments. Teachers responded to questions by indicating their preferences on a scale. If they had additional comments they were encourated to place them in the space so provided.;

2. Open-ended questions with ample space provided for responses.;
3. Questions requiring the teacher to fill in or supply some limited information e.g., numbers of students participating, types of students that experienced difficulty with the module, time spent on a particular task, etc.

The instruments in the Log used by teachers to monitor or spot problems as they occurred during the participation phase of the module was similar in form to the Lesson Evaluation Form developed for use in the Comprehensive Career Education Project, the school based model. (This applies only for the first 3 modules tested. For the last 4 modules tested, the mid and post module questionnaires were used in place of the logs.) This particular questionnaire if extensively filled out by teachers and if extensively corroborated by other data sources will be a most valuable input into the module revision process.

3. Approximate Length

To begin, the length of the Teacher Evaluation Log, in terms of both teacher time and pages, is dependent on the length and complexity of the module. It was assumed that during the entire life span (3 1/2 - 6 weeks) of a module the Log would take a minimum of about 2 1/2 hours to complete; at a maximum it would expand to nearly 5 hours. As noted earlier the Log attempts to elicit open-ended responses from teachers and hence a great deal of paper space is required. Checklists, while possible, would have necessitated some preliminary observations of module classroom behavior before the instruments could have been meaningfully constructed. Secondly, the time frame for instrument development would have gone beyond that available. So the LOG's are large in appearance which may, in turn, have affected the

extent to which they were completed by participating teachers.

The two questionnaires (and the optional forms) used for testing 4 modules in the Spring of 1974 were quite a bit shorter than the LOG. The two questionnaires required approximately 25-30 minutes apiece to complete.

4. Scoring Procedures

Since only a maximum of 4 teachers were included in the experimental groups, the scoring of the LOG'S is fairly simple. Scaled as well as open-ended responses will be summed and collated respectively. No detailed scoring is anticipated at this time. (Questionnaires used for the 4 modules tested in the Spring of 1974 will be scored in a manner similar to that described for the LOG'S above.)

E. Panel Reviews Procedure

1. General Structure

The panel reviews were designed to collect in-depth information for revision. After completion of a module, teachers were convened in a panel setting to discuss, at length, each part of a module. They were informed of the panel review session prior to their starting of a module. For the sessions they were asked to bring in their LOGS, questionnaires and/or any other notes they might have made.

The panel session was not simply for the purpose of duplicating the results obtained from the LOG or questionnaires but specifically it was to probe for more in-depth opinions of teachers. The panel leader(s) was (were) responsible for generating a written report summarizing

the comments of teachers. It should also be noted that in the session, the teachers were given the opportunity to express their feelings as to how the module should be revised. This type of input from classroom practitioners is frequently invaluable for revision purposes.

For the 4 modules tested in the Spring of 1974 two panel reviews involving teachers were conducted. The first occurred midway through the module and the second, at the end of the module. The general structure described above was employed for both panel reviews.

2. Types of Questions

Basically, only 4 questions were used for the panel review. They are:

- What do you feel were the strengths or strong points in this part of the module?;
- What do you feel were the weaknesses of this part of the module?;
- If you did try to overcome the weakness, what specifically did you do and how well did it work?;
- What recommendations for change would you make in this part of the module?

The panel leader was instructed to probe into these questions as teachers responded. For example, why was this a strength, how did your students respond, why was this a weakness, did this part require too much time, etc. The panel leader probed also for the degree of agreement amongst the teachers. In addition to their perception of module parts, teachers were asked about their feelings toward the entire module. Notes were kept by either the panel leader or an observer if one was present.

3. Approximate Length

Length in this case is dependent on many factors - length of the

module, module complexity, module problems, interest of the teachers, etc. As a general rule, at least 2 1/2 hours was required when a single panel was conducted for a module. (When two were conducted somewhat more than one hour was required per session.) Frequently, for the single panel case, more than one session was needed.

4. Scoring Procedures

For revision purposes, ne panel review report will simply be summarized and included in the Reviser's Information Summary.

F. Observer Form (Modules tested in the Spring of 1974, only)

1. General Structure

For the 4 modules tested in the Spring of 1974, classroom observation was also utilized as a means of collecting data. At each school involved in the Spring testing, an observer was stationed to provide an independent source of information regarding how each module was being implemented. A very brief three page observation form was provided for the observers to complete as they watched students and teachers. It was felt that data of this type would be a valuable source of input into the revision process.

To facilitate both the observation process and the use of audio-visual devices the four simulations were conducted in a special room designated for that purpose at each school. The situation was somewhat akin to a laboratory setting.

2. Type of Questions

There are three basic types of questions included in the Observer Form. The first is concerned with the use of booklets and media during the class period. The second involves comments on observations the observer has made during the period but which do not seem to fit into any

category on the form. The third major category relates to group or individual activities in the simulation as well as to the interaction between students and teachers. This type of information will serve as an anchor for other types collected in the evaluation.

3. Approximate Length

The form is short (3 pages) and designed to be used relatively easily in the classroom by the observer. Directions as well as written examples were provided to facilitate completion of the observer forms.

4. Scoring Procedures

To the extent feasible the data collected from the observers will be compared and if possible collated. To a degree, this was a trial run at this method of obtaining information. Comparative study may indicate that it cannot be collated and that for future project use the forms and the procedure may have to become more standard in nature.

V. SUMMARY - Further Considerations

As the reader will recall, the front section of the evaluation notebook was designed to give prospective users of the notebook an overview of both the Occupational Exploration Project and the evaluation that was conducted during the project year 1973-1974. Due to the heavy revision/formative focus of the first year evaluation, the text of the front matter contains frequent references to the revision process. It is hoped that this will help to facilitate the reviser's job.

The succeeding sections of the notebook are organized into module specific sections. These start out with a brief introduction or description of the module that was evaluated, followed by the nature of the actual evaluation that was carred out. Data tables and

interpretation come directly after the textual discussion. Last, but of greatest importance to the reviser, is the Reviser's Information Summary (RIS). This is the summary of the evaluation information as it relates to revision. The reviser should read both the introductory text and the RIS in the module specific section first. If he/she desires more detailed information, the data tables may be referred to.

It should be emphasized as a closing note that the information in the RIS does not constitute a prescription or a mandatory set of rules for the reviser to follow. The RIS is a summary of the information collected and interpretation of trends and their potential meaning for revision. The evaluation data collected this year could not begin to anticipate nor answer all revision questions. Only a limited subset of the questions could be answered in this evaluation. Hence, the judgment of the reviser will still play a major role in the revision process. That judgment, however, should be tempered against curriculum criteria and the revision plan or strategy should contain a clear rationale as to why changes were undertaken.